6-9-2020

Hot and Cold Seasons in the Housing Market: Comment

Dean Scrimgeour
Colgate University, dscrimgeour@colgate.edu

Hot and Cold Seasons in the Housing Market: Comment

Dean Scrimgeour*

Colgate University

June 9, 2020

Abstract

Using American Housing Survey data from 1999 only, Ngai and Tenreyro (2014) show households who move in the summer occupy their home for longer and have fewer and less costly renovations soon after purchase, pointing to superior match quality during the thicker summer market. However, applying the same methods to other years of the American Housing Survey eliminates or substantially weakens these results. Furthermore, Ngai and Tenreyro's result on duration of occupancy is driven in part by the particular way Ngai and Tenreyro measure duration in years, rather than months.

JEL Codes: R21, R31
Keywords: housing, seasonality, duration models

*Contact: Dean Scrimgeour, Economics Department, Colgate University, 13 Oak Drive, Hamilton, NY 13346. Email: dscrimgeour@colgate.edu. Thanks to Emily Conover, Brian Mendez-Smith, Chad Sparber and seminar participants at Colgate University for suggestions and advice.

1. Introduction

Ngai and Tenreyro (2014) interpret seasonal patterns in housing markets using a matching model in which the thicker markets of the hot summer period allow buyers acquire a house closer to their ideal.[1][2] Ngai and Tenreyro present evidence that households who move during the summer stay in their house longer, undertake fewer repairs and spend less on repairs. In particular, using data from the American Housing Survey (AHS) in 1999, Ngai and Tenreyro estimate that summer movers occupy their residences 5% longer (standard error 2%), and that within two years of moving in they undertake 12% fewer replacements or additions (standard error 5%), and spend 18% less as a share of the house's value (standard error 7%).

In this paper, I show that Ngai and Tenreyro's conclusions are overstated. I estimate Ngai and Tenreyro's regressions using data from the 2001 through 2011 waves of the AHS in addition to 1999. These are all the waves that ask recent movers enough about their prior move to allow measurement of a completed spell occupying a single residence. I include both National Survey and Metropolitan Survey data. Across the three outcomes Ngai and Tenreyro study, the 1999 wave usually generates the largest estimated effect when using any single year's data. When pooling all available years and using methods essentially identical to Ngai and Tenryro, I estimate that summer movers stay in their homes 2.4% longer (standard error 0.7%), and make 1.6% fewer replacements and additions (standard error 1.7%) and spend 5% less on replacements and additions shortly after moving. The magnitude of these estimated effects is about one sixth to one half of what Ngai and Tenreyro report.

Apart from the discrepancy across years of the AHS, another source of doubt about the results in Ngai and Tenreyro's paper arises from how they measure occupancy duration in years, rather than months. For example, a household that moves in December 1997 and again in January 1998 is said to have occupied their home for one year, the same as another household that moves first in January 1997 and again in December 1998. When estimating their regressions with the

[1] A summer move is defined as one occurring between 1 April and 30 September.

[2] See also Novy-Marx (2009) for a broader discussion of the differences between hot and cold housing markets.

duration outcome measured in months, the estimated difference between summer and winter moves is less than half of what Ngai and Tenreyro find and is no longer statistically significant.

In summary, Ngai and Tenreyro's estimates are substantially weakened when a broader sample of data is used and when a more fine-grained measure of the occupancy duration is used. While there are undoubtedly major seasonal variations in housing markets, and matches might be better in hotter periods, Ngai and Tenreyro overstate the evidence that match quality is superior for those who move home in the high season.

2. Further Waves of the American Housing Survey

Ngai and Tenreyro use data only from the 1999 wave of the AHS. This was the first wave to ask about the residents' move prior to the move that brought them to their current address. Questions about the prior move also featured on subsequent waves of the AHS up until 2011. In this section I show that most of the key estimates in Ngai and Tenreyro (2014) are substantially smaller in magnitude and no longer statistically significant in the later waves of the survey.[3] It appears that the magnitude and statistical significance of Ngai and Tenreyro's estimates are driven by the use of 1999 data only.

2.1. Occupancy Duration

Ngai and Tenreyro find that occupancy lasts longer when a household moves during the summer. The estimated effect appears anomalous due to the sample used. In particular, the 1999 wave of the AHS suggests larger contrasts between summer and winter movers than most other years in which the survey runs.

While Ngai and Tenreyro present estimates from a variety of specifications, these different specifications do not much affect their key estimates. Therefore, I

[3] I use data both from the national survey and the metropolitan survey. The national survey is conducted biennially, so I use data from six more waves from the national survey. I use data from the metropolitan survey from 2002, 2004, 2007 and 2009. Starting in 2011 the national and metropolitan surveys were combined. The prior move questions are asked only of households who have moved in the prior two years. For 2007 and 2009, most of the data come from the national sample. For 2011 most of the data are from the metropolitan sample.

present results from a single specification.[4] The specification here corresponds to column (5) in Ngai and Tenreyro's Table 4. Figure 1 shows point estimates and 95% confidence intervals for the coefficient on summer across different years of the survey. Estimates from 1999 data that Ngai and Tenreyro used are in red, and estimates from other years in black. (Multiplying by 100 gives the estimated percentage difference in duration between summer and winter movers.) Regression tables corresponding to these estimates are available in an online appendix.

Whereas Ngai and Tenreyro estimate that households who move in the summer stay 5.1% longer based on 1999 surveys, Figure 1 shows that most other waves of the survey produce results that are smaller in magnitude and not statistically significant. The 2009 data generates slightly larger estimates (5.4%) than the 1999 wave, but other years' coefficients are smaller, and in 2005 negative. When pooling all AHS data, allowing for year fixed effects, I estimate that summer movers stay about 2.4% longer, with a standard error of 0.7%.

2.2. Number of Repairs and Additions

Ngai and Tenreyro also use the number of repairs and additions made in the two years following purchase to measure match quality. They show that winter moves are associated with more renovations, implying a poorer match. However, when running the same regressions on other waves of the survey, this effect is smaller in magnitude, occasionally opposed in sign, and not statistically significant. See Figure 2, where the red bar uses 1999 data as in Ngai and Tenreyro. Ngai and Tenreyro estimate 11.5% fewer repairs for summer purchases using the

[4] I also make minor changes in the data used, relative to Ngai and Tenreyro. I only use households that report income, even if income is not included in the regression, and I exclude households who report a prior move that occurs after the current move, which is obviously a data error. I also impute some variables, such as heating and cooling degree days or Consolidated Metropolitan Statistical Area using information from more narrowly-defined Standard Metropolitan Statistical Area. These changes or imputations do not affect the essential conclusions I present. As in Ngai and Tenreyro, I weight the observations using the AHS weights. In comparing my tables with Ngai and Tenryro, note that there appear to be several transcription errors in Ngai and Tenreyro (2014)'s Table 4. The replication files imply that columns (3) through (7) all have urban status (METRO3) fixed effects. Columns (3) and (4)'s estimates should be shifted right one column, while column (5)'s are a duplication of column(4). Furthermore, column (6) should indicate heating/cooling degree day controls. None of these errors has an important effect on the reported coefficients, especially on the summer move variable.

1999 wave. Pooling all years of the survey I find a difference of 1.6% (with a standard error of 1.7%).

2.3. Cost of Renovations

The third measure of match quality Ngai and Tenreyro use is the cost of repairs and additions, relative to the house's value. For summer matches, they find repairs made within two years of moving to be about 17.8% less than winter matches. As shown in Figure 3, this is much larger than implied by other waves. When pooling all years together, I estimate an effect of 5.0%. Given a standard error of 2.7%, the estimate is not statistically significant at the five percent level, and is certainly not consistent with the 18% difference between summer and winter movers that Ngai and Tenreyro find.

3. Measuring Duration in Months Instead of Years

I now present a distinct concern about the estimates in Ngai and Tenreyro. They measure the duration of occupancy in years, ignoring the month of the transactions, despite using this same month of transaction to determine whether the move was a summer or winter move. I present regression results using occupancy duration measured in months and show that this greatly reduces the estimated difference between summer and winter moves. I also discuss whether this difference could be interpreted as a bias and review the underlying data on reported month of move.

3.1. Comparison of Estimates using Years and Months of Occupancy

The blue bars with open squares in Figure 1 show confidence intervals for the difference in match duration between summer and winter moves across different years of the AHS. The underlying regressions for the blue confidence intervals use duration in months as the outcome variable. Contrast these with the black bars, which use the Tenreyro-Ngai duration measure.

Ngai and Tenreyro's estimates are larger than I obtain using duration in month for the 1999 data. This discrepancy is consistent across years – the estimated difference between summer and winter moves is always more negative when using duration in months. The estimates using duration in months are never statistically significant at the five percent level. These conclusions are not sensitive to the particular specification that Ngai and Tenreyro use. Adding heating and cooling degree days or CMSA fixed effects, for example, do not alter the conclusions.

Measuring duration in years coarsens the outcome variable, implicitly rounding the underlying duration in a way that has a different effect, on average, for summer and winter moves, and so affects the estimates.[5] Take the rounded component of duration as

$$12 \times duration_y - duration_m$$

where $duration_y$ is duration in years and $duration_m$ is duration in months. When regressing the rounded component of $duration_y$ on the summer dummy, the coefficient is about 0.5: duration in years is more likely to have been rounded up for summer moves and down for winter moves. Converting durations into years as Ngai and Tenreyro do adds about half a month more to the duration for summer movers than for winter movers.

An alternative method of measuring durations in whole years simply rounds to the nearest year. Anything up to five months is a zero-year-duration, from six to seventeen months is a one-year-duration, and so on. While the estimate for 1999 from Ngai and Tenreyro is about a 5% difference between summer and winter movers, using the rounding approach, the difference is about 0.5% (standard error 0.7%), no longer significant. This is smaller, but not statistically significantly different from, the estimates obtained using duration in months as the outcome. See Table A.7.

3.2. Monte Carlo Evidence

I now consider a simulation that explains the divergence between estimates using years and those using months duration for the outcome variable. In each of these

[5]Dempster and Rubin (1983) and Heitjan and Rubin (1991) discuss regression analysis with rounded data.

exercises the true data generating process features no difference between summer and winter in terms of subsequent duration. Nevertheless, Ngai and Tenreyro's procedure is capable of generating a nonzero estimated effect on average.

Consider a simulation in which the month of the prior move is taken as given in the data, then the occupancy duration is sampled independently (with replacement) from the empirical distribution of reported durations (in months).[6] For each such bootstrap sample we can compute the duration in years (as in Ngai and Tenreyro) and estimate the a bivariate poisson regression using either months or years as the measure of duration.

Figure 4 shows the results of these simulations. While the distribution of $\hat{\beta}$ is centered on zero when I use duration in months, it is substantially higher on average when using years. This difference is 0.022, on average, and is one estimate of the bias from measuring the duration in years. (It is also very close to the 0.024 difference between the corresponding estimates from the actual data in 1999.) The scatter plot shows that the coefficient when using the duration in years is always greater than when using months. From the histogram, we can derive a bootstrapped p-value for an observed $\hat{\beta}$ at the observed level, and this is 0.128, larger than the 0.023 computed without regard to the biases documented here. (This is a two-sided p-value, measuring the fraction of samples in which the estimated coefficient is larger than the corresponding estimate from the actual data.)

4. Conclusion

Ngai and Tenreyro explore an aspect of search and matching markets – that match quality is improved in thick markets. They present evidence for this idea using American Housing Survey data from 1999. I have shown in this paper that their evidence is fragile. First, data from other years of the survey show much smaller, and statistically insignificant, differences between summer and winter movers in terms of all three dimensions Ngai and Tenreyro explore: occupancy

[6]An alternative approach, which shows quantitatively similar results,samples the duration in months from a geometric distribution calibrated to have the same average duration as in the 1999 data, while the initial month is taken from the data, as before.

duration, the number of repairs and additions and the cost of such repairs and additions. I find estimated differences across seasons between one seventh and one quarter of what was previously reported.

Furthermore, Ngai and Tenreyro's estimator for the difference in occupancy duration uses a measure of occupancy duration that seems to generate an upward bias in the estimated effect of summer moves. Using a finer measure of occupancy duration, in months instead of years, or rounding the months of occupancy to the nearest number of completed years, further weakens their result.

References

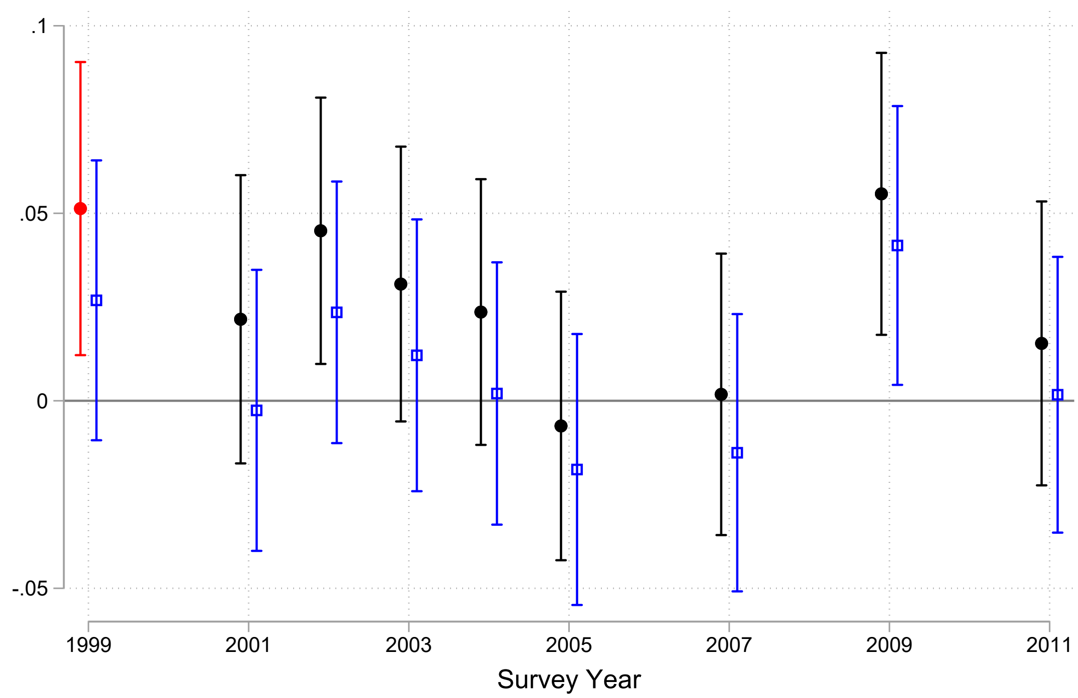
Dempster, Arthur P. and Donald B. Rubin, "Rounding error in regression: The appropriateness of Sheppard's corrections," *Journal of the Royal Statistical Society: Series B (Methodological)*, 1983, *45* (1), 51–59.

Heitjan, Daniel F. and Donald B. Rubin, "Ignorability and Coarse Data," *The Annals of Statistics*, 1991, *19* (4), 2244–2253.

Ngai, L. Rachel and Silvana Tenreyro, "Hot and cold seasons in the housing market," *American Economic Review*, 2014, *104* (12), 3991–4026.

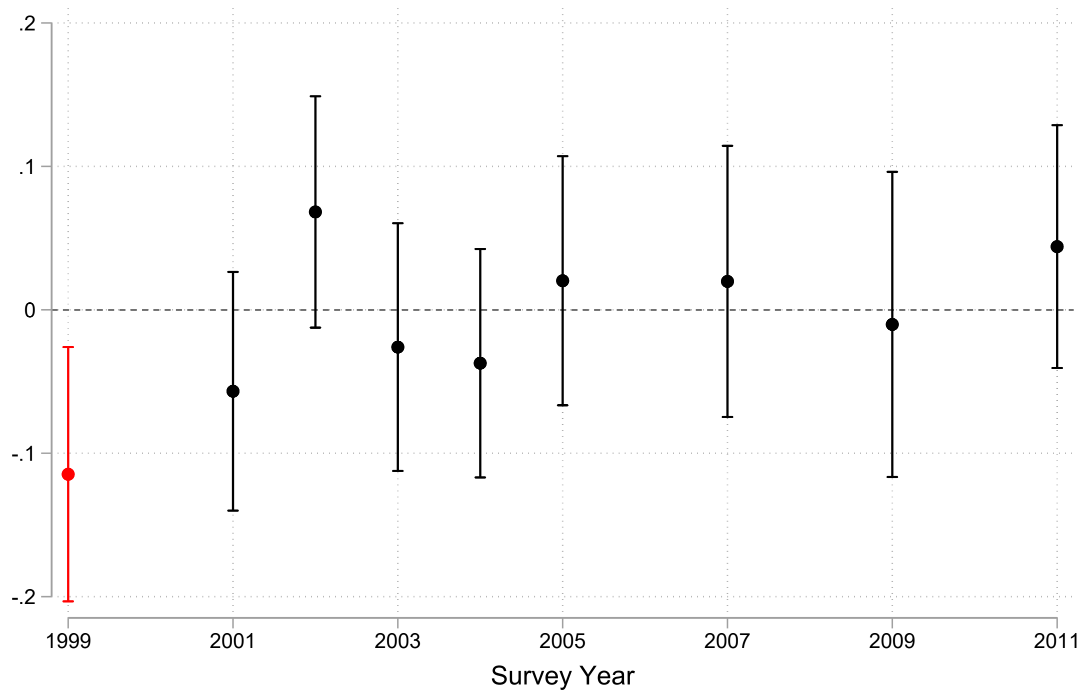
Novy-Marx, Robert, "Hot and cold markets," *Real Estate Economics*, 2009, *37* (1), 1–22.

Figure 1: Duration of the Match and Season in which Match was Formed, across AHS survey years



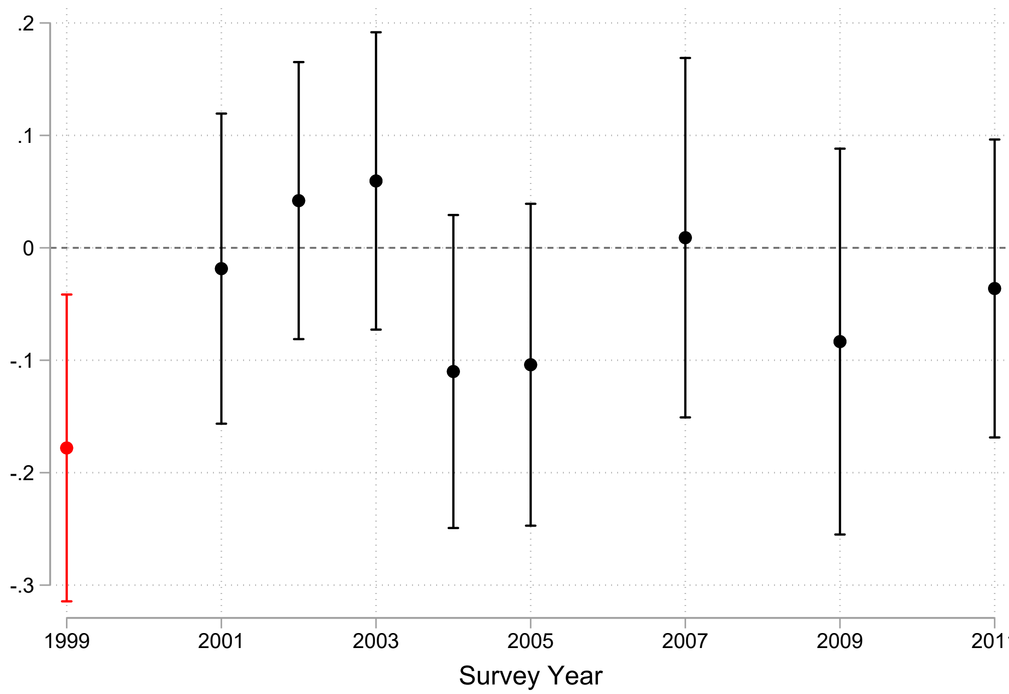
Notes: the figure plots confidence intervals for the parameter on summer, interpreted as one-hundredth of the percentage difference in duration between summer moves and winter moves, for different years of the AHS survey and for different outcomes. The red bar shows estimates corresponding to Ngai and Tenreyro's, while black bars show estimates from the same specification using other years. The blue bars show estimates that use duration measured in months as the outcome. The estimates are taken from regressions corresponding to column (5) in Table 4 of Ngai and Tenreyro (2014). In addition to summer, these regression include as covariates log family income, the number of people in the household, the number of adults in the household, and region fixed effects. See the appendix for regression tables.

Figure 2: Number of Replacements and Additions and Season in which Match was Formed, across AHS survey years



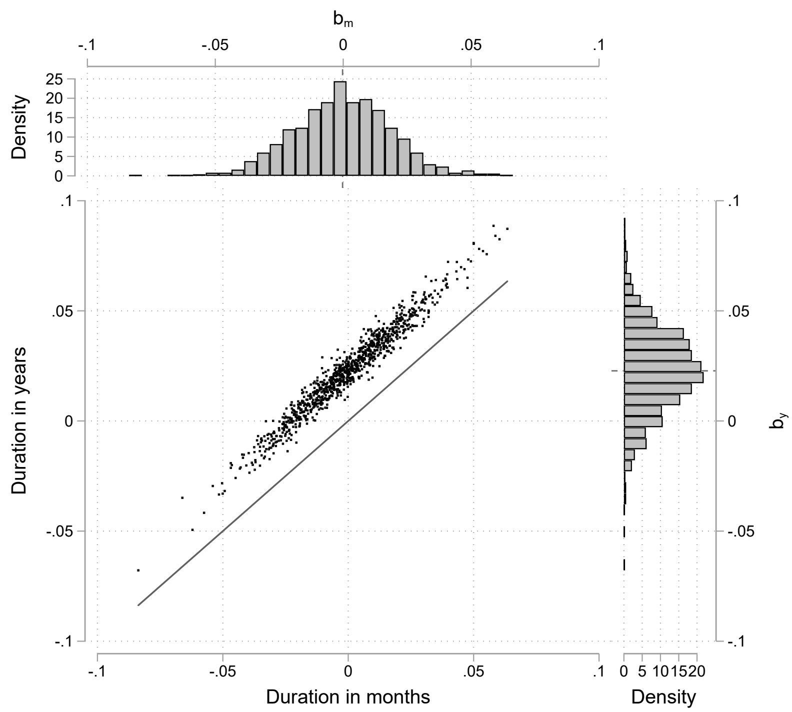
Notes: the figure plots confidence intervals for the parameter on summer, interpreted as one-hundredth of the percentage difference in the number of replacements and additions between summer moves and winter moves, for different years of the AHS survey. The red bar is from 1999, the year Ngai and Tenreyro used. The estimates are taken from regressions corresponding to column (5) in Table 5 of Ngai and Tenreyro (2014). In addition to summer, these regression include as covariates log of the year the housing unit was built, log family income, the number of people in the household, the number of adults in the household, and region fixed effects. See the appendix for regression tables.

Figure 3: Costs of Replacements and Additions and Season in which Match was Formed, across AHS survey years



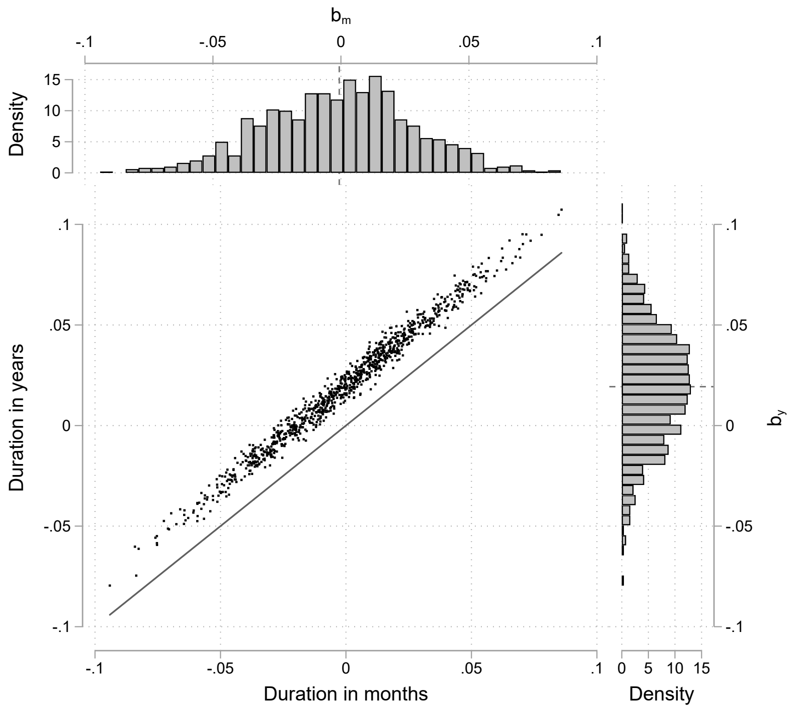
Notes: the figure plots confidence intervals for the parameter on summer, interpreted as one-hundredth of the percentage difference in the cost of replacements and additions (as a share of the housing unit's value) between summer moves and winter moves, for different years of the AHS survey. The red bar is from 1999, the year Ngai and Tenreyro used. The estimates are taken from regressions corresponding to column (5) in Table 6 of Ngai and Tenreyro (2014). In addition to summer, these regression include as covariates log of the year the housing unit was built, log family income, the number of people in the household, the number of adults in the household, and region fixed effects. See the appendix for regression tables.

Figure 4: Monte Carlo: sampling months duration from the empirical distribution



Notes: the figure plots estimates of the coefficient on summer in regressions that use either duration in years (coefficient is b_y) or duration in months (coefficient is b_m) as the outcome, with summer move as the explanatory variable. There is a joint distribution, represented by the scatter plot, as well as marginal distributions. The estimates come from 1,000 different simulations. The simulations sample from the empirical distribution of duration in months from 1999 data.

Figure 5: Monte Carlo: sampling months duration from a geometric distribution



Notes: the figure plots estimates of the coefficient on summer in regressions that use either duration in years (coefficient is b_y) or duration in months (coefficient is b_m) as the outcome, with summer move as the explanatory variable. There is a joint distribution, represented by the scatter plot, as well as marginal distributions. The estimates come from 1,000 different simulations. The simulations sample durations in months from a geometric distribution calibrated to have the same mean as the empirical distribution from 1999 data.

FOR ONLINE PUBLICATION

A Regression Tables

A1. Regression Estimates Using Different Survey Years

Tables A.1-A.4 present estimates from regressions underlying Figure 1-3. They correspond to column 5 in each of Tables 4 - 6 of Ngai and Tenreyro (2014).

Table A.1: Duration Regressions (AHS, each wave separately, using duration in years)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Summer	0.0513	0.0217	0.0453	0.0311	0.0237	-0.00671	0.00172	0.0552	0.0153	0.0238
	(0.0199)	(0.0196)	(0.0181)	(0.0187)	(0.0181)	(0.0183)	(0.0191)	(0.0192)	(0.0193)	(0.00706)
(log) Income	0.0937	0.102	0.0660	0.0739	0.0735	0.0681	0.0785	0.0739	0.0847	0.0814
	(0.0119)	(0.0128)	(0.0102)	(0.0104)	(0.0102)	(0.0111)	(0.00990)	(0.00948)	(0.00933)	(0.00397)
HH size	0.0613	0.0396	0.0319	0.0461	0.0475	0.0423	0.0245	0.0359	0.0238	0.0390
	(0.00827)	(0.00892)	(0.00779)	(0.00812)	(0.00734)	(0.00798)	(0.00821)	(0.00768)	(0.00824)	(0.00303)
Num. adults	-0.0942	-0.0848	-0.0556	-0.114	-0.0837	-0.0881	-0.0502	-0.0679	-0.0354	-0.0755
	(0.0176)	(0.0174)	(0.0156)	(0.0175)	(0.0165)	(0.0177)	(0.0161)	(0.0161)	(0.0156)	(0.00621)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban status fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	No	No	No	No	No	Yes
Observations	6688	5760	8780	5876	6470	6047	6737	6310	12173	64841
Survey Year	1999	2001	2002	2003	2004	2005	2007	2009	2011	all

Notes: Poisson regression. Summer is a dummy variable, equal to one if the household head's previous move was in the spring or summer (April to September). Sample includes all respondents who report a nonnegative duration measured in months and household income. Income is household income, HH Size is the number of people in the household, while Num. adults is the number of adults in the household. Urban status categories are (i) central city of MSA; (ii) inside MSA, but not in central city-urban; (iii) inside MSA, but not in central city-rural; (iv) outside MSA, urban; (v) outside MSA, rural. Robust standard errors in parentheses.

Table A.2: Duration Regressions (AHS, each wave separately, using duration in months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Summer	0.0268	-0.00255	0.0236	0.0121	0.00194	-0.0183	-0.0139	0.0414	0.00161	0.00624
	(0.0190)	(0.0191)	(0.0178)	(0.0185)	(0.0178)	(0.0184)	(0.0189)	(0.0190)	(0.0188)	(0.00693)
(log) Income	0.0862	0.101	0.0668	0.0804	0.0788	0.0742	0.0848	0.0826	0.0831	0.0842
	(0.0117)	(0.0130)	(0.0104)	(0.0106)	(0.0102)	(0.0111)	(0.00987)	(0.00948)	(0.00898)	(0.00397)
HH size	0.0598	0.0377	0.0337	0.0460	0.0450	0.0416	0.0274	0.0373	0.0242	0.0391
	(0.00784)	(0.00865)	(0.00754)	(0.00803)	(0.00730)	(0.00798)	(0.00808)	(0.00756)	(0.00804)	(0.00297)
Num. adults	-0.0885	-0.0786	-0.0569	-0.116	-0.0790	-0.0838	-0.0504	-0.0716	-0.0349	-0.0740
	(0.0170)	(0.0173)	(0.0154)	(0.0172)	(0.0165)	(0.0178)	(0.0159)	(0.0157)	(0.0153)	(0.00611)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban status fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	No	No	No	No	No	Yes
Observations	6688	5760	8780	5876	6470	6047	6737	6310	12173	64841
Survey Year	1999	2001	2002	2003	2004	2005	2007	2009	2011	all

Notes: Poisson regression. Summer is a dummy variable, equal to one if the household head's previous move was in the spring or summer (April to September). Sample includes all respondents who report a nonnegative duration measured in months and household income. Income is household income, HH Size is the number of people in the household, while Num. adults is the number of adults in the household. Urban status categories are (i) central city of MSA; (ii) inside MSA, but not in central city-urban; (iii) inside MSA, but not in central city-rural; (iv) outside MSA, urban; (v) outside MSA, rural. Robust standard errors in parentheses.

Table A.3: Replacements and Additions (AHS, each wave separately)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Summer	-0.115	-0.0568	0.0682	-0.0260	-0.0372	0.0203	0.0198	-0.0102	0.0441	-0.0160
	(0.0452)	(0.0425)	(0.0411)	(0.0441)	(0.0406)	(0.0443)	(0.0482)	(0.0543)	(0.0432)	(0.0170)
(log) Income	0.184	0.129	0.237	0.160	0.0949	0.148	0.190	0.148	0.0829	0.149
	(0.0326)	(0.0255)	(0.0321)	(0.0302)	(0.0253)	(0.0292)	(0.0354)	(0.0328)	(0.0293)	(0.0114)
HH size	0.0562	0.0423	0.0241	0.0181	0.0242	-0.0346	-0.00287	0.0432	-0.0169	0.0165
	(0.0177)	(0.0172)	(0.0164)	(0.0185)	(0.0168)	(0.0192)	(0.0200)	(0.0213)	(0.0201)	(0.00702)
Num. adults	-0.0544	0.00434	-0.0277	0.0221	0.0502	0.0864	-0.0179	0.0176	0.136	0.0235
	(0.0352)	(0.0334)	(0.0343)	(0.0347)	(0.0335)	(0.0368)	(0.0338)	(0.0429)	(0.0346)	(0.0132)
(log) Year built	-29.13	-23.71	-38.22	-24.23	-30.23	-28.73	-29.47	-30.23	-21.77	-26.94
	(1.997)	(1.708)	(1.645)	(1.483)	(1.441)	(1.526)	(1.692)	(1.767)	(1.495)	(0.624)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	No	No	No	No	No	Yes
Observations	5883	5468	6473	5603	5677	5383	5865	4137	9972	54461
Survey Year	1999	2001	2002	2003	2004	2005	2007	2009	2011	all

Notes: Poisson regression. The outcome is the number of replacements and additions. Summer is a dummy variable, equal to one if the household head's move into the current residence was in the spring or summer (April to September). Sample includes all respondents who report a nonnegative duration measured in months and household income. Income is household income, HH Size is the number of people in the household, while Num. adults is the number of adults in the household. Year build is the year the housing unit was built. Robust standard errors in parentheses.

Table A.4: Cost of Replacements and Additions (AHS, each wave separately)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Summer	-0.178	-0.0185	0.0420	0.0595	-0.110	-0.104	0.00901	-0.0834	-0.0362	-0.0497
	(0.0696)	(0.0704)	(0.0629)	(0.0675)	(0.0710)	(0.0730)	(0.0816)	(0.0875)	(0.0676)	(0.0272)
(log) Income	-0.0560	-0.0706	0.00644	-0.0535	0.00849	-0.0274	-0.0446	-0.118	-0.0974	-0.0635
	(0.0404)	(0.0357)	(0.0386)	(0.0368)	(0.0498)	(0.0482)	(0.0504)	(0.0556)	(0.0366)	(0.0158)
HH size	-0.0474	-0.0316	-0.0662	-0.0772	-0.0500	-0.102	-0.0875	-0.0697	-0.0519	-0.0651
	(0.0293)	(0.0335)	(0.0263)	(0.0305)	(0.0319)	(0.0319)	(0.0338)	(0.0362)	(0.0333)	(0.0121)
Num. adults	-0.0112	0.0185	-0.0571	0.0589	-0.0176	0.124	-0.0155	0.0713	0.0804	0.0436
	(0.0601)	(0.0650)	(0.0550)	(0.0723)	(0.0717)	(0.0668)	(0.0711)	(0.0866)	(0.0649)	(0.0255)
(log) Year built	-16.16	-8.153	-21.01	-7.110	-18.15	-6.638	-8.636	-14.13	-18.50	-11.46
	(2.882)	(2.791)	(2.990)	(2.662)	(2.853)	(2.914)	(3.181)	(3.183)	(2.557)	(1.074)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	No	No	No	No	No	Yes
Observations	3014	2848	3174	2756	2885	2552	2985	2092	5966	28272
Survey Year	1999	2001	2002	2003	2004	2005	2007	2009	2011	all

Notes: The outcome is the log of the cost of replacements and additions relative to the value of the housing unit. Summer is a dummy variable, equal to one if the household head's move into the current residence was in the spring or summer (April to September). Sample includes all respondents who report a nonnegative duration measured in months and household income. Income is household income, HH Size is the number of people in the household, while Num. adults is the number of adults in the household. Year build is the year the housing unit was built. Robust standard errors in parentheses.

A2. Duration in Years or Months

Tables A.5 and A.6 present regression estimates of the effect of a summer move on occupancy duration using Ngai and Tenreyro (2014)'s measure of duration in years then the duration in months. Each table uses all AHS data available, subject to the restriction that the sample used is the same across specifications. The table demonstrates that specification changes are relatively unimportant for the estimates, in contrast to the choice of outcome variable. These correspond to Table 4 in Ngai and Tenreyro (2014), though the last two columns are omitted because changes in the heating/cooling degree days and CMSA variables over time make these last columns more difficult to estimate with all years' data.

Table A.5: Duration of the Match and Season in which Match was Formed, all AHS survey years (Ngai and Tenreyro outcome)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Summer	0.0201	0.0198	0.0205	0.0215	0.0238	0.0235	0.0236
	(0.00713)	(0.00706)	(0.00706)	(0.00706)	(0.00706)	(0.00706)	(0.00706)
(log) Income		0.0817	0.0773	0.0752	0.0814	0.0809	0.0808
		(0.00384)	(0.00389)	(0.00389)	(0.00397)	(0.00396)	(0.00398)
HH size				0.0162	0.0390	0.0392	0.0391
				(0.00242)	(0.00303)	(0.00304)	(0.00304)
Num. adults					-0.0755	-0.0758	-0.0758
					(0.00621)	(0.00620)	(0.00620)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban status fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Degree Days	No	No	No	No	No	Yes	Yes
CMSA fixed effects	No	No	No	No	No	No	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	64841	64841	64841	64841	64841	64841	64841

Notes: Poisson regression. The dependent variable is the duration of stay in years. Summer is a dummy variable, equal to one if the household head's previous move was in the spring or summer (April to September). Sample includes all respondents who report a nonnegative duration measured in months and household income. Income is household income, HH Size is the number of people in the household, while Num. adults is the number of adults in the household. Urban status categories are (i) central city of MSA; (ii) inside MSA, but not in central city-urban; (iii) inside MSA, but not in central city-rural; (iv) outside MSA, urban; (v) outside MSA, rural. Heating/cooling degree days categories are (i) Coldest: 7,001+ heating degree days and $<$ 2,000 cooling degree days; (ii) Cold: 5,500–7,000 heating degree days and $<$ 2,000 cooling degree days; (iii) Cool: 4,000–5,499 heating degree days and $<$ 2,000 cooling degree days; (iv) Mild: $<$ 4,000 heating degree days and $<$ 2,000 cooling degree days; (v) Mixed: 2,000–3,999 heating degree days and 2,000+ cooling degree days; (vi) Hot: $<$ 2,000 heating degree days and 2,000+ cooling degree days. Robust standard errors in parentheses.

Table A.6: Duration of the Match and Season in which Match was Formed, all AHS survey years (Months Duration)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Summer	0.00261	0.00228	0.00289	0.00392	0.00624	0.00609	0.00637
	(0.00701)	(0.00693)	(0.00694)	(0.00693)	(0.00693)	(0.00693)	(0.00693)
(log) Income		0.0845	0.0803	0.0782	0.0842	0.0838	0.0839
		(0.00383)	(0.00389)	(0.00389)	(0.00397)	(0.00396)	(0.00398)
HH size				0.0167	0.0391	0.0392	0.0391
				(0.00238)	(0.00297)	(0.00297)	(0.00297)
Num. adults					-0.0740	-0.0743	-0.0744
					(0.00611)	(0.00611)	(0.00610)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban status fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Degree Days	No	No	No	No	No	Yes	Yes
CMSA fixed effects	No	No	No	No	No	No	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	64841	64841	64841	64841	64841	64841	64841

Notes: Poisson regression. The dependent variable is the duration of stay in months. Summer is a dummy variable, equal to one if the household head's previous move was in the spring or summer (April to September). Sample includes all respondents who report a nonnegative duration measured in months and household income. Income is household income, HH Size is the number of people in the household, while Num. adults is the number of adults in the household. Urban status categories are (i) central city of MSA; (ii) inside MSA, but not in central city-urban; (iii) inside MSA, but not in central city-rural; (iv) outside MSA, urban; (v) outside MSA, rural. Heating/cooling degree days categories are (i) Coldest: 7,001+ heating degree days and < 2,000 cooling degree days; (ii) Cold: 5,500–7,000 heating degree days and < 2,000 cooling degree days; (iii) Cool: 4,000–5,499 heating degree days and < 2,000 cooling degree days; (iv) Mild: < 4,000 heating degree days and < 2,000 cooling degree days; (v) Mixed: 2,000–3,999 heating degree days and 2,000+ cooling degree days; (vi) Hot: < 2,000 heating degree days and 2,000+ cooling degree days. Robust standard errors in parentheses.

A3. Alternate Measure of Duration in Years

Table A.7 presents regression estimates corresponding to Table 4 in Ngai and Tenreyro (2014), but using a different measure of duration in years, rounding months to the nearest multiple of twelve.

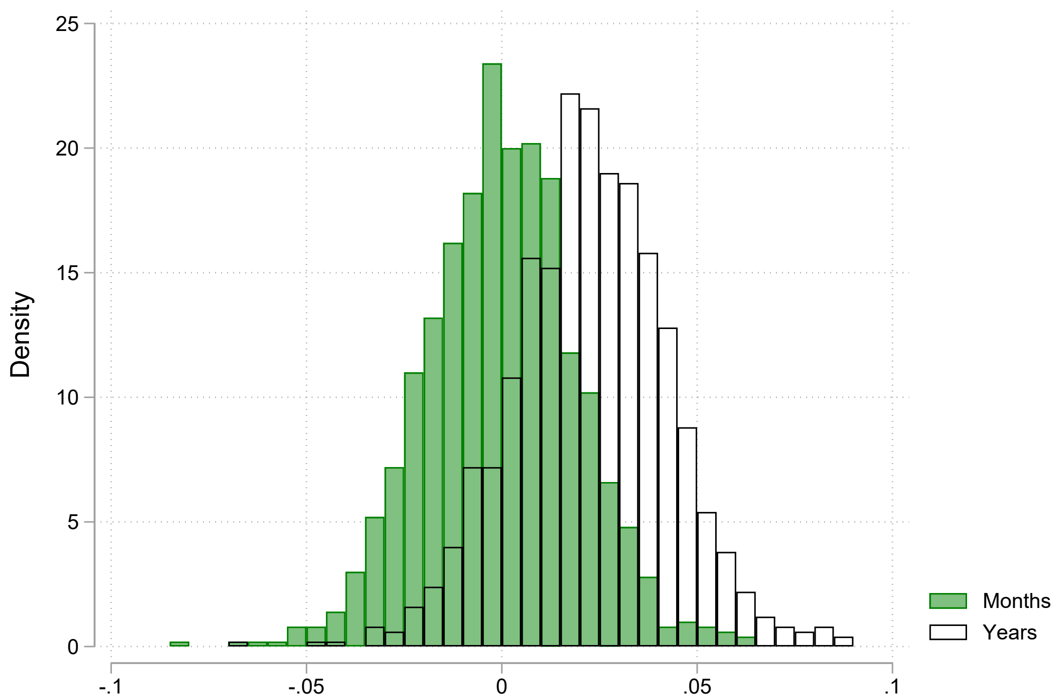
Table A.7: Duration of the Match and Season in which Match was Formed (duration rounded to the nearest year)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Summer	0.00183	0.00150	0.00215	0.00317	0.00549	0.00535	0.00562
	(0.00707)	(0.00699)	(0.00700)	(0.00699)	(0.00699)	(0.00699)	(0.00699)
(log) Income		0.0848	0.0804	0.0783	0.0844	0.0839	0.0841
		(0.00385)	(0.00391)	(0.00391)	(0.00399)	(0.00398)	(0.00400)
HH size				0.0165	0.0389	0.0390	0.0389
				(0.00240)	(0.00299)	(0.00300)	(0.00300)
Num. adults					-0.0741	-0.0743	-0.0744
					(0.00617)	(0.00617)	(0.00617)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban status fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Degree Days	No	No	No	No	No	Yes	Yes
CMSA fixed effects	No	No	No	No	No	No	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	64841	64841	64841	64841	64841	64841	64841

Notes: Poisson regression. Outcome is duration in years, having rounded the duration in months to the nearest year. Robust standard errors in parentheses. See notes to Table A.5 for more details.

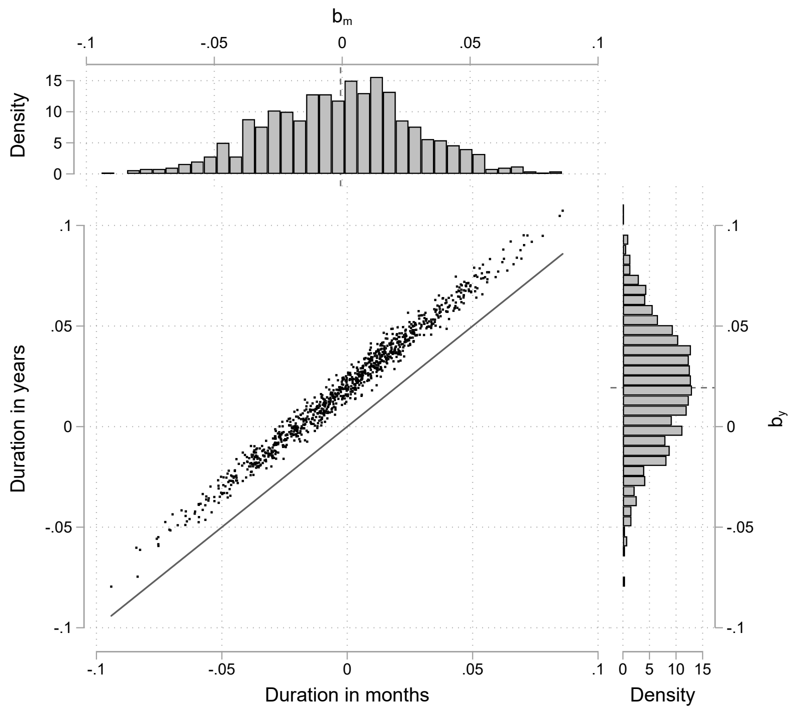
B Additional Graphs

Figure B.1: Monte Carlo: sampling months duration from the empirical distribution



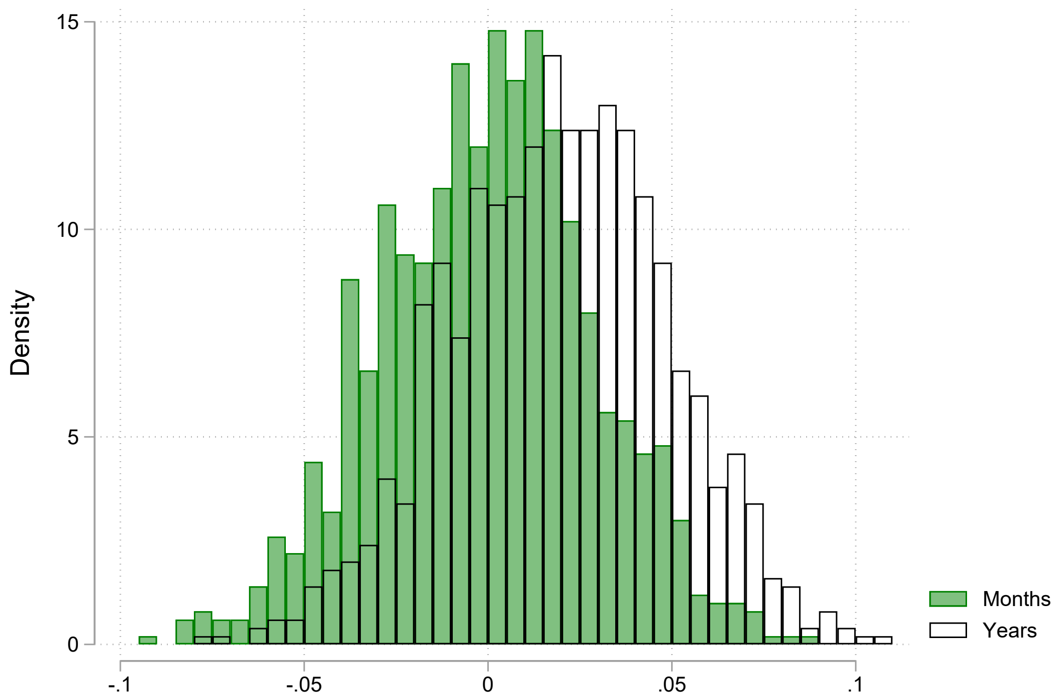
Note: These two histograms are from the simulation described in the main text (see Section 3.2.). The simulations draw at random from the distribution of measured durations (in months) for each observation in the bootstrap sample. The duration in years is then calculated from the month of prior move together with the duration in months. The distribution of the coefficient on summer move is centered on zero when the outcome is duration in months but is biased upward when the outcome is duration in years.

Figure B.2: Monte Carlo: sampling months duration from a geometric distribution



Note: The scatter plot and histograms are from a simulation that follows the one previously described except that it samples durations from a geometric distribution that has the same mean as the empirical distribution. The duration in years is then calculated from the month of prior move together with the duration in months. The distribution of the coefficient on summer move is centered on zero when the outcome is duration in months but is biased upward when the outcome is duration in years.

Figure B.3: Monte Carlo: sampling months duration from a geometric distribution



Note: These two histograms are from the simulation whose outcomes are also shown in Figure B.2. The simulations draw the duration in months at random from a geometric distribution that has the same mean as the empirical distribution. The duration in years is then calculated from the month of prior move together with the duration in months. The distribution of the coefficient on summer move is centered on zero when the outcome is duration in months but is biased upward when the outcome is duration in years.